Precision Livestock Farming '17

Edited by D. Berckmans and A. Keita

Papers presented at the 8th European Conference on Precision Livestock Farming

Nantes, France
12-14 September '17

Traditional mixed linear modelling versus modern machine learning to estimate cow individual feed intake

C. Kamphuis, J.W. van Riel, R.F. Veerkamp, and R.M. de Mol
Wageningen University & Research, Livestock Research, PO Box 338, the Netherlands
Claudia.Kamphuis@wur.nl

Abstract

Three modelling approaches were used to estimate cow individual feed intake (**FI**) using feeding trial data from a research farm, including weekly recordings of milk production and composition, live-weight, parity, and total FI. Additionally, weather data (temperature, humidity) were retrieved from the Dutch National Weather Service (KNMI). The 2014 data (245 cows; 277 parities) were used for model development. The first model (**M1**) applied an existing formula to estimate energy requirement using parity, fat and protein corrected milk, and live-weight, and assumed this requirement to be equal to energy intake and thus FI. The second model used 'traditional' Mixed Linear Regression, first using the same variables as in M1 as fixed effects (**MLR1**), and then by adding weather data (**MLR2**). The third model applied Boosted Regression Tree, a 'modern' machine learning technique, again once with the same variables as M1 (**BRT1**), and once with weather information added (**BRT2**). All models were validated on 2015 data (155 cows; 165 parities) using correlation between estimated and actual FI to evaluate performance. Both MLRs had very high correlations (0.91) between actual and estimated FI on 2014 data, much higher than 0.46 for M1, and 0.73 for both BRTs. When validated on 2015 data, correlations dropped to 0.71 for MLR1 and 0.72 for MLR2, and increased to 0.71 for M1 and 0.76 for both BRTs. FI estimated by BRT1 was, on average, 0.35kg less (range: -7.61 – 13.32kg) than actual FI compared to 0.52kg less (range: -11.67 – 19.87kg) for M1. Adding weather data did not improve FI estimations.

Keywords: precision feeding, dairy cows, Big Data, prediction, machine learning

Introduction

Feed efficiency in dairy cattle is gaining interest due to the limited availability of natural resources (De Mol et al., 2016) and the challenge to feed over 9 billion people by 2050 (FAO, 2016). Feed efficiency is a measure on the efficient conversion of feed intake into milk production (Lu et al., 2015). Feed intake (**FI**) can be measured using dedicated feeding equipment, e.g., roughage intake control systems, but these systems are exclusively used under experimental conditions. Actual daily feed intake of dairy cows under commercial circumstances, thus, remains unknown.

During the past decades farm sizes have increased, milk yields have risen and automation and sensor technologies for milking and other farm tasks (e.g., observing cows in oestrus) have become increasingly popular on dairy farms (Mottram, 2016). This increase in automation and sensor technology has also increased the availability of the amount and sources of data. A shift from traditional methods to analyse data to new modelling approaches is, therefore, expected to occur (Mottram, 2016). One of these new modelling approaches are Big Data analytics. The mainstream definition of Big Data involves three Vs (Sonka and Cheng, 2015): (1) Volume, or the amount of available data (2) Velocity, or the speed and frequency of data arrival and processing (Devlin, 2012; Zaslavsky et al., 2012) and the capability to respond to events (near) real-time, and (3) Variety, or the availability of different (un)structured formats of data, e.g., spreadsheets, drone images, and pictures. Over time, other Vs have been added like variability, veracity (trustworthiness of data), visualisation, and value (McNulty, 2014). The tools used in Big Data analytics include data-driven techniques like machine learning.

To explore the usefulness of Big Data analytics for the estimation of cow individual FI, this study compared three different modelling approaches for estimating FI. The first model estimated energy requirements based on parity, live-weight, and milk production, using the assumption that energy requirement equals energy intake. The second model used traditional Mixed Linear Regression to predict feed intake, and the third model was a machine learning algorithm called Boosted Regression Tree. Since Variety is one of the three characteristics used in the mainstream definition of Big Data, the latter two modelling approaches were repeated with weather information as additional data source.

Materials and Methods

<u>Experimental data</u>

Data were retrieved from 10 feeding trials conducted in 2014 and/or 2015 at the Wageningen University & Research farm in Lelystad. During each trial, FI (kg dry matter /day) was recorded on a daily basis as the sum of roughage intake (kg dry matter /day) and concentrate (kg /day) intake, where the latter was converted into kg dry matter by multiplying the intake with 0.89. This value was considered the Gold Standard and further referred to as ‘actual FI’. For each calendar week, daily cow individual FI recordings were converted into an average FI. The same approach was used to calculate an average daily live-weight per calendar week. Data on milk yield and composition were recorded once a week. From these data fat and protein corrected milk (**FPCM**) was calculated using (CVB, 2008; Klop et al., 2016):

$$FPCM\ (kg/day) = Milk\ yield * (0.337 + fat\% * 0.116 + protein\% * 0.06), \quad (1)$$

where milk yield is the recorded milk yield expressed in kg per day, fat% and protein% are the percentages of fat and protein, respectively. Averages of actual FI and live-weight of the same week that milk yield and composition were recorded were added. Lastly, for each cow, parity and the week in milk were added for each calendar week.

In addition to FI and cow information, weather data from Lelystad Airport (~10km distance from the research farm) were retrieved from the Royal Dutch Meteorological Institute (KNMI). This institute records freely accessible weather variables on a daily basis. From all available variables, the minimum, maximum, and average temperature and relative humidity per day were selected. Based on these recorded variables, the daily temperature (and relative humidity) range, and the difference between todays average temperature (and relative humidity) with the average temperature (and relative humidity) over the past seven days were derived. Weather information was then coupled with FI and cow information based on date, and thus, all cows had once-a-week data recordings of FPCM, live-weight, parity, DIM, and weather information available.

Cows with less than four records (that is, less than one month of data) were excluded from further analysis. Live-weight recordings and FCPM values that were outside the mean ± four times the standard deviation were considered outliers and set at missing. Cow-weeks that had missing values for any of the recorded or derived parameters were excluded (28.3%). These exclusion criteria resulted in 407 cows and 3,787 cow-weeks from seven feeding trials available for further analysis. Table 1 summarizes the number of cows, cow-weeks, and treatments per trial, as well as the range in parity per trial, the range in DIM per trial, and the year in which the trial was conducted. Cows could be included in multiple trials, and in trials crossing years. Therefore, the number of unique cows in this study (n = 300) was lower than the 407 cows reported in Table 1. The number of unique cows in 2014 was 245; in 2015 there were 155 unique cows.

Table 1: Characteristics of trials used for training and testing including the number of cows, cow-days, and treatments (Treatm.) per trial (Trial), the range of parity and days in milk (DIM) per trial, and the year(s) in which the trial was conducted.

Trial	Treatm. (n)	Cows (n)	Cow-days (n)	Parity range	DIM range	Year
1	4	136	399	1 – 7	70 – 160	2014
2	3	52	75	1 – 6	42 – 207	2014–2015
3	3	96	2,594	1 – 7	7 – 364	2014–2015
4	1	10	65	3 – 5	6 – 56	2014
5	1	15	39	1 – 5	21 – 56	2014
6	3	39	177	2 – 9	14 – 140	2015
7	5	59	438	1 – 5	7 – 63	2014–2015
Total	20	407	3,787			

Statistical analysis

Several models (Table 2) were developed using data from 2014. Each of these models was tested on the same data used for training, as well as on new, independent, data from 2015. The first model (**M1**; Table 2) used an existing model to estimate energy requirement, and assumed this requirement to be equal to energy intake. Energy requirement was calculated using the Dutch net energy evaluation for dairy cows (Van Es, 1975; CVB, 2008):

$$VEM/day = (42.4 * LW^{0.75}) + (442 * FPCM)) * (1 + (FPCM - 15) * 0.00165) \quad (2)$$

Where LW represents a cow's live-weight (in kg), and FPCM refers to the fat and protein corrected milk (in kg; formula 1). Since further details on used feeding and treatments within these trials were unknown, we assumed that provided feed had 975 VEM/kg dry matter. Therefore, required VEM/day (from formula 2) was divided by 975 to compute FI / day.

The second model used the Mixed Linear Regression approach (Table 2). The first variant of this model (**MLR1**) used FI as dependent variable, and the same variables used as M1 as fixed effects, where parity was included as a three-level factor (parity 1, 2, and ≥3). Trial, treatment within trial, cow, week in milk, and month of the year were included as random effects. The second variant of the Mixed Linear Regression model (**MLR2**) extended the MLR1 with weather information by adding temperature and relative humidity data as fixed effects.

The third model used machine learning to estimate FI (Table 2), using a nonlinear predictive method called Regression Tree (James et al., 2015). It involves segmenting the predictor space using binary splits into smaller regions that contain training observations that are similar. Typically, the mean of all training observations falling into such a small region is used as predicted outcome for a new observation (not used for training) that belongs to that same region. The power of trees lies in the simple method, and visualising the tree makes the model itself easy to interpret. However, single trees are often large and over-fitted, and consequently lack predicting accuracy on new, unseen observations. Improving the predictive performance of trees is possible, e.g., by aggregating many trees (James et al., 2015). Boosting is such an approach to generate many trees and aggregate them into one single outcome. Boosting creates multiple small trees sequentially, where each new tree uses the residuals from the previous tree as response (James et al., 2015). The first variant of this third model applied Boosted Regression Tree (**BRT1**) using FI as independent variable, and the same variables as M1, where parity was included as a three-level factor (1,2, and ≥3) variable, and with week in milk and month of the year added to the model. The second variant of this third model (**BRT2**) extended BRT1 with weather information by adding temperature and relative humidity data. Both ensemble trees consisted of 1,000 sub-trees with each sub-tree having a maximum number of four splits.

Table 2: short description and variables included per model.

Model	Description	Variables
M1	Energy requirement according to formula 1 and feed intake according to formula 2	parity, live-weight, fat and protein corrected milk
MLR 1	Mixed Linear Regression without weather info	Fixed effects: * Random effects: trial, treatment within trial, cowid, week in milk, month of the year
MLR 2	Mixed Linear Regression with weather info	Fixed effects: *, temperature[1], humidity[2] Random effects: trial, treatment within trial, cowid, week in milk, month of the year
BRT1	Boosted Regression Tree without weather info	*, week in milk, month of the year
BRT2	Boosted Regression Tree with weather info	*, temperature[1], humidity[2], week in milk, month of the year

* same variables as listed for model M1; 1 includes average temperature of the past week, and the absolute difference between today's temperature and the average temperature of the past week; 2 includes average humidity of the past week, and the absolute difference between today's humidity and the average humidity of the past week

To evaluate performance of each model in predicting FI, the Pearson's correlation between predicted FI and actual FI was calculated for each model for both the training set (2014 data), and the test set (2015 data). For both MLR1 and MLR2 (Table 2) only coefficients of the fixed effects were used to predict FI. Additionally, the mean difference between predicted and actual FI was calculated for the test set only. This was done for all observations combined, per parity category (1, 2, and ≥3), and per lactation stage (<100, 100-200, and >200 days in milk).
All analyses were conducted using RStudio (using R version 3.1.1; R Core Team 2016; James et al., 2015) extended with the following packages: RODBC (Ripley and Lapsley, 2016), plyr (Wickham, 2011), lme4 (Bates et al., 2015), Hmisc (Harrell, 2016), data.table (Dowle et al., 2015), and gbm (Ridgeway, 2015).

Results

The average actual FI of cows was 21.2kg for both the training (2014) and test data (2015). Table 3 summarizes the correlations between actual and estimated FI by the different models, for both the training (2014) and the test (2015) data. Both MLR models have high correlations between actual and estimated FI for the training set, indicating a good fit. Correlations for M1 and both BRT models on the training data were lower. When models were applied to observations not used for training, correlations between actual and estimated FI dropped for both MLR models. In contrast, correlations for M1 and both BRT models increased. Correlations were similar between models with and without weather information, regardless whether training or test data were used. All models had comparable correlations when applied on the test set, and all models estimated, on average, FI to be lower than actual FI. Estimated FI from MLR2 deviated most, on average, from the actual FI. Although M1 had a low mean difference (-0.52, Table 3), it did have the highest range in difference between actual and estimated FI; estimated FI ranged to be almost 12kg less than actual FI to almost 20kg too much. The range in difference between actual and estimated FI was lowest with ~20kg for both MLR models.

Table 3. Per model the correlation between estimated feed intake (FI) and actual FI on the training (2014) and on the test (2015) data, and the mean and range of the difference (both in kg) between the estimated and actual FI on the test data (2015).

Model	Training set	Test set		
	Correlation	Correlation	Mean difference (kg)	Range difference (kg)
M1	0.46	0.71	-0.52	-11.67 – 19.87
MLR1	0.91	0.71	-1.23	-7.70 – 12.32
MLR2	0.91	0.72	-1.73	-8.24 – 11.76
BRT1	0.73	0.76	-0.35	-7.61 – 13.32
BRT2	0.73	0.76	-0.35	-7.61 – 13.32

Table 4 summarizes correlations for different parity categories and lactation stages. Both BRT models have high correlations between actual and estimated FI for first parity cows, in contrast to the MLR models. Also, BRT models had the highest correlations for cows earlier in lactation, whereas correlation dropped for both BRT models for cows later in lactation. The M1 and both MLR models have the highest correlation for cows that are 100 to 200 days in lactation. Again, there is no difference in correlation between models that do not include weather information (MLR1, BRT1) versus those that had this information included (MRL2, BRT2).

Table 4. Per model the correlation between estimated feed intake (FI) and actual FI using the test data (2015) per parity group, and per category of days in milk (DIM). The number of records per category of parity or DIM is listed between brackets

Model	Parity			DIM		
	1 (98)	2 (461)	≥3 (831)	<100 (544)	100-200 (299)	>200 (547)
M1	0.67	0.77	0.71	0.60	0.77	0.63
MLR1	0.67	0.76	0.71	0.58	0.75	0.64
MLR2	0.66	0.76	0.71	0.58	0.75	0.65
BRT1	0.82	0.79	0.72	0.78	0.69	0.38
BRT2	0.82	0.79	0.72	0.78	0.69	0.38

Discussion

The current study is not the first one estimating cow individual FI using machine learning. Van der Waaij et al. (2016) analysed a dataset very similar to the one we used, with the exception that they had additional sensor information (rumination) and a differentiation between roughage and concentrate intake. But there are three more differences with that study worthy to discuss: firstly, they included cow identification as proxy for the influence of genotype on FI and reported a positive influence of this variable in predicting FI. In contrast, we left cow identification out of the equation, since BRT will likely use that variable as root node which will likely result in improved cow individual FI prediction. However, generalization to new data (that is, unseen cow identification numbers) will not work since the model will not recognize this new 'value'. Secondly, adding weather information did not contribute to a better FI prediction in our study, whereas Van der Waaij et al. (2016) reported temperature having a 'positive influence'. Unfortunately, they did not specify the magnitude of that positive contribution nor provided results of a model without weather information, leaving the question unanswered whether temperature adds significantly to FI prediction. Thirdly, Van der Waaij et al. (2016) used a Neural Network, and reported this network to be unable to predict FI in case of missing data. Given that sensor data are incomplete by definition, Neural Networks may not be the appropriate analytical tool to be used in practice.

The majority of the data used in the current study (68.5%) originated from a single feeding trial crossing years (Experiment 3, Table 1). Thus, data used for training (2014) were not independent from data used for testing (2015) which may have overestimated results. Still, both MLRs and BRTs were trained and tested on same data, and thus, results are relative to each other. The MLR models performed well on the training set, but correlations dropped substantially when applied on the test set, indicating a possible overfit of these regression models. In contrast, BRT appeared to be more robust and less prone to overfitting, since correlations on the test set were similar to those of the training set. All models underestimated actual FI, with the M1 having the widest range in differences between actual and estimated FI (Table 3). Also large differences in correlations between parity groups and lactation stages were seen, for all three modelling approaches. BRT models appear to estimate FI for first parity cows and those early in lactation much better than M1 and MLR models, whereas MLR and M1 seem to outperform BRT for cows later in lactation (Table 4). Future research should investigate whether the differences between model performance on the test set are significant, why all models consistently underestimate actual FI, and what is causing the differences between modelling approaches for different parity groups or lactation stages.

The data used in the current study were pre-processed such that all three modelling approaches could handle the data (e.g., records with missing values for any of the recorded or derived variables were excluded to allow mixed linear regression analysis). By doing so, we ensured training and testing of different models to be conducted on the same data, but we may have limited the potential of the BRT models in two ways: firstly, we excluded almost 30% the records with missing data for one or more predictor variables, whereas these incomplete records may still hold potentially valuable information. Machine learning approaches like BRT are known for their capability to deal with these incomplete data. Secondly, by excluding almost 30% of the data we reduced the volume considerably, whereas machine learning works requires large amounts of data. Saying this, even if we would have included all data, critics could argue that even then we would not have enough data for machine learning and that we have linked this study to Big Data incorrectly. On the other side, the volume characteristic is highly subjective, depending on the industry and application, and a specific threshold on this characteristic is lacking (Sonka and Cheng, 2015). In the near future, data of the current study will be extended with 15 years of feeding trials, conducted at several different research farms, and including additional data sources, like breeding values for FI, roughage and concentrate percentages fed, and other sensor data. This will certainly increase volume, but will also add complexity due to increased variety and velocity of the available data.

Conclusion

Three modelling approaches were used to estimate cow individual FI. The ‘traditional’ MLR models had high correlations on the training data, but these dropped substantially when the models were applied to the test data. In contrast, the ‘modern’ BRT models had lower correlations on the training data, but appeared to be more robust since correlations remained similar on the test set. Moreover, FI estimated by BRT1 was, on average 0.35kg less than actual FI, compared with the commonly applied M1 model that had an average predicted FI more than 0.5kg less than actual FI. Adding weather information did not improve FI estimations. To better meet the three Vs of Big Data, and potentially improve performance of machine learning algorithms that thrive on large volumes of data, future research will focus on including more farms, more feed experiments conducted during more years, and adding data from more sources.

Acknowledgements

This project is part of the project KB-27-001-001 Big Data for healthy resources utilization, which is funded by the Dutch Ministry of Economic Affairs. Colleagues at Wageningen University & Research are acknowledged for sharing the data.

References

Bates, D., M. Mächler, B. Bolker, and S. Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. 2015 67(1):48.

CVB. 2008. Table Booklet feeding of ruminants.

De Mol, R. M., R. M. A. Goselink, J. W. Van Riel, H. M. Knijn, and A. T. M. Van Knegsel. 2016. The relation between eating time and feed intake of dairy cows. Pages 387-392 in Precision Dairy Farming 2016. C. Kamphuis and W. Steeneveld, ed. Wageningen Academic Publishers, Wageningen, the Netherlands, Leeuwarden, the Netherlands.

Devlin, B. 2012. The Big Data Zoo - Taming the beasts: the need for an integrated platform for enterprise information. 9sight Consulting. from http://9sight.com/pdfs/Big_Data_Zoo.pdf

Dowle, M., A. Srinivasan, T. Short, and S. Lianoglou, 2015. data.table: Extension of Data.frame. from https://CRAN.R-project.org/package=data.table

FAO, 2016. Retrieved 13 December 2016 from http://www.fao.org/livestock-environment/en/.

Harrell, F. E., 2016. Hmisc: Harrell Miscellaneous. from https://CRAN.R-project.org/package=Hmisc.

James, G., D. Witten, T. Hastie, and R. Tibshirani. 2015. An introduction to statistical learning with applications in R. Springer Texts in Statistics. Springer New York, Heidelberg, Dordrecht, London.

Klop, G., B. Hatew, A. Bannink, and J. Dijkstra. 2016. Feeding nitrate and docosahexaenoic acid affects enteric methane production and milk fatty acid composition in lactating dairy cows. Journal of Dairy Science 99(2):1161-1172.

Lu, Y., M. J. Vandehaar, D. M. Spurlock, K. A. Weigel, L. E. Armentano, C. R. Staples, E. E. Connor, Z. Wang, N. M. Bello, and R. J. Tempelman. 2015. An alternative approach to modeling genetic merit of feed efficiency in dairy cattle. Journal of Dairy Science 98(9):6535-6551.

McNulty, E., 2014. Understanding Big Data: the seven V's. Retrieved 24 March, 2017 from http://dataconomy.com/2014/2005/seven-vs-big-data/.

Mottram, T. 2016. Animal board invited review: Precision livestock farming for dairy cows with a focus on oestrus detection. Animal 10(10):1575-1584.

Ridgeway, G. 2015. gbm: Generalized Boosted Regression Models. from https://cran.r-project.org/web/packages/gbm/gbm.pdf.

Ripley, B. and M. Lapsley, 2016. ODBC connectivity. from https://cran.r-project.org/web/packages/RODBC/vignettes/RODBC.pdf.

Sonka, S. and Y.-T. Cheng. 2015. Big Data: More than a lot of numbers! Department of Agricultural and Consumer Economics, University of Illinois Urbana-Champaign 5: 201.
Van der Waaij, B. D., R. L. Feldbrugge, and R. F. Veerkamp. 2016. Cow feed intake prediction with machine learning. Pages 377-382 in Precision Dairy Farming 2016. C. Kamphuis and W. Steeneveld, ed. Wageningen Academic Publishers, Wageningen, the Netherlands, Leeuwarden, the Netherlands.
Van Es, A. J. H. 1975. Feed evaluation for dairy cows. Livestock Production Science 2(2):95-107.
Wickham, H. 2011. The Split-Apply-Combine Strategy for Data Analysis. 2011 40(1):29.
Zaslavsky, A., C. Perera, and D. Georgakopoulos. 2012. Sensing as a servive and big data. Pages 21-29 in Proc. International Conference on Advances in Cloud Computing (ACC-2012).